CPB Netherlands Bureau for
Economic Policy Analysis

Attrition and sample selection in the pre-COOL and COOL studies

This background document describes the pre-COOL and COOL cohort studies, using linked CBS micro-data. The micro-data allows to compare the children from the two cohort studies with all children in the Netherlands of the same age (pre-COOL) or in the same school grade (COOL). We discuss the representativeness of the samples, as well as panel attrition within the cohort studies.

The pre-COOL and COOL cohort studies provide a rich data set on skills and background of children in the Netherlands. Pre-COOL surveys children before they enter primary school, COOL follows children in primary and secondary education.

CPB Background Document

Maria Zumbuehl

December 2020

1 Overview

The notitie "Ongelijkheid van het jonge kind" documents and discusses differences in achievement across various (socio-economic) groups in the Netherlands. The empirical results are based on data from the pre-COOL and COOL studies, Dutch cohort studies that together survey children from the ages of two to fifteen, and that have been collected between 2007 and 2014.

This background document provides insight into the datasets. We link the COOL and pre-COOL cohort studies to CBS microdata. On behalf of OCW and in agreement with Kohnstamm, who created the datasets, we use register data to examine the representativeness of the cohort samples, as well as panel attrition and reasons for dropping out.

The pre-COOL study follows one cohort of young children from age two to six, with yearly surveys between 2010 and 2014. The sample is not designed to be representative for the Netherlands. Children from families with lower socio-economic status are oversampled. We provide population weights to adjust for varying probabilities to be part of the study. The study follows the same children from the start cohort until age six, which allows for a panel dataset. However, we do find evidence for selective panel attrition.

The COOL study surveys school classes in intervals of three years, with measurements in 2007, 2010 and 2013. The COOL sample is designed as a representative sample of the population of Dutch school children. The sample of children in primary education is additionally augmented by adding more children from disadvantaged backgrounds, which allows better inference for specific groups of special interest. Comparing the full sample (core and augmented) with register data, we find that in primary education lower socio-economic groups are indeed oversampled, turning to an oversampling of higher socio-economics groups in secondary education. Following a class (and not student) centered approach, COOL is not well suited for panel studies, as students who have to repeat a class are dropped from the sample. There is also significant attrition and subsequent resampling of schools that participate in the study, which decreases the size of a potential panel cohort.

We find small, but not overwhelming differences in the distribution of achievement measures when comparing them with and without using population weights. This is reassuring for previous studies that have not been able to augment the data with information from CBS microdata. While the population data are particularly important when adding the different waves of the pre-COOL and COOL study together, within age groups the bias of not using these weights is not very large and the results are likely to hold.

2 Pre-COOL

2.1 Data in brief

The pre-COOL study surveys children before they enter primary school. It consist of two cohorts (starting at age two and at age four) that are followed over time. In this document, as well as in the notitie "Ongelijkheid van het jonge kind", we focus on the larger cohort that samples children at age two. The study started in 2010, gathering data on skills, behavior and background characteristics of two-year-old children at daycare institutions (instellingscohort) or at their home (gezinscohort). In 2011 the tests were repeated, with the now three-year-old children of the first wave and a sample of other three-year-old children who had newly started preschool by then. These two first waves constitute the basis for the further panel.[1]

2.2 Link with CBS register data

The match between the pre-COOL sample and the CBS microdata is not complete. Based on names and birthdates, the CBS was able to find a match in the register data for the majority, but not all of the children from the pre-COOL study. That means that all of the following tests and analyses, as well as the sample weights will not be based on the complete pre-COOL data, but rather on the matched subset. While the population weights we construct take care of the representability issue, a direct comparison to other studies that are using only pre-COOL data has to be done with caution.

From all measurement waves 12.7% of all observations cannot be matched to the register data. This number is even slightly higher if we look at only the two waves which form the basis of the panel, with 13.9% unmatched observations in the first wave and 15.0% in the second wave respectively. Most of the unsuccessful matches come from the instellingscohort, where the total success rate is 82.8% (compared to the gezinscohort with a match of 98.7%).[2]

Using the background data and tests scores that are in the pre-COOL data, we see that the probability of a match to the CBS data is not random. Children who cannot be linked to the register data are more likely to come from a more disadvantaged socio-ethnic background (lower parental education and non-western migration background). While the information on the socio-ethnic background is not available for all children, other information, such as the language test scores are available for most of them. We find that the not matched children score significantly lower on the early language tests than children whom we can link to the register data. Using population weights can help to alleviate the problem, but only if the not matched group does not additionally differ in important unobservable characteristics.

2.3 Sample selection

2.3.1 Representativeness

Does the pre-COOL sample represent the relevant age group within the Netherlands? We compare the pre-COOL sample (in wave 1 or 2) to the reference population of 2 or 3 year-olds in the Netherlands, in the

[1] A detailed description of the study design and surveys can be found on http://www.pre-cool.nl/.
[2] The gezinscohort is based on a representative sample of individual children that was provided by the CBS. For the instellingscohort day care centers were sampled and the children of the participating day care centers were looked up in the register data subsequently. The different sampling approach explains the differences in matching success.

corresponding test period (2010 and 2011). The pre-COOL group consists out all those pre-COOL children of each wave, who have been matched to the CBS (see above). The reference population comprises all children who were born in 2008 between April and October, which is the sample that the pre-COOL children were drawn from. The group of the reference population also includes the pre-COOL children.

Table 1 The pre-COOL sample differs from the reference population in key variables, such as migration background

VARIABLES	Wave 1		Wave 2	
	Reference population	Pre-COOL	Reference population	Pre-COOL
Female	0.487	0.491	0.487	0.498
Migration: Native	0.748	0.682	0.743	0.666
Migration: Western	0.076	0.072	0.080	0.072
Migration: non-western second generation	0.168	0.242	0.167	0.255
Migration: non-western first generation	0.008	0.005	0.010	0.007
Parental education: unknown	0.133	0.088	0.138	0.098
Parental education: max. mbo2	0.198	0.214	0.196	0.222
Parental education: mbo3-4/havo/vwo	0.259	0.271	0.258	0.278
Parental education: hbo/wo	0.409	0.426	0.408	0.401
Household income quintile: unknown	0.030	0.006	0.035	0.005
Household income quintile: lowest	0.213	0.241	0.203	0.246
Household income quintile: second	0.202	0.211	0.200	0.212
Household income quintile: middle	0.231	0.227	0.232	0.241
Household income quintile: fourth	0.187	0.191	0.188	0.175
Household income quintile: highest	0.138	0.124	0.142	0.121
Type of household: unknown	0.037	0.010	0.041	0.010
Type of household: couple with children (married)	0.270	0.295	0.246	0.271
Type of household: couple with children (married)	0.618	0.600	0.626	0.625
Type of household: single parent	0.075	0.095	0.086	0.094
Number of obs.	114091	2182	114979	2406

In both waves children with a migration background (more specifically second generation non-western) are oversampled as can be seen in Table 1. We further see that the pre-COOL sample in both waves contains more children from families in lower income quintiles, as well as more children who are living in a single-parent household. These patterns hold also for the later waves (three to five).

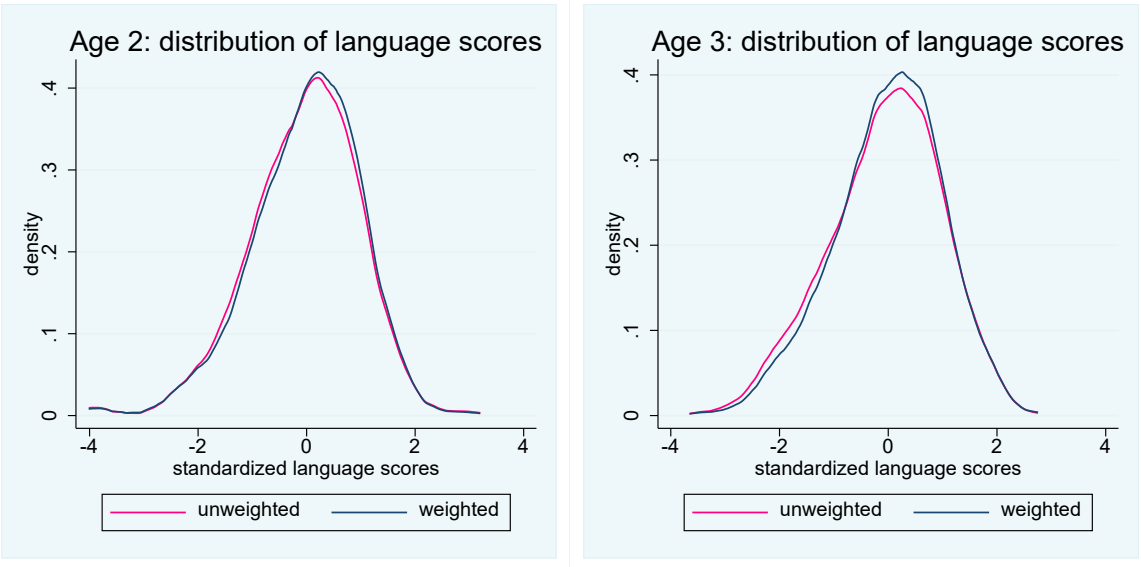
2.3.2 Population weights

We generate population weights that allow us to account for the fact that not all children are equally likely to be sampled in the pre-COOL study. Specifically, for each measurement wave of pre-COOL we construct population weights, that capture the inverse probability of each observation in this specific wave to be in the sample. Since the sample composition differs between waves, individuals can have different weights in different waves.

We use the complete reference population (including the pre-COOL sample) to estimate the probability to be in the pre-COOL sample, based on background characteristics, such as migration background, parental income or education. The weights are constructed as the inverse of the predicted probability that we estimate with a probit regression.

Reweighing the sample with population weights can have an impact on the distribution of the outcome variable. Figure 1 displays an example of how the outcome variable shifts, if the oversampled variables are correlated with the outcome variable. It shows the distribution of language skills for a weighted and an unweighted sample in waves 1 and 2 respectively. While the differences are small, we can see that if we reweigh the sample to give underrepresented groups a higher weight, the distribution of the language skills shifts to the right, and the average becomes higher. We find no such difference for the work/play attitude of the children.

Figure 1 The weighted pre-COOL sample displays marginally higher language test scores than the unweighted sample



2.4 Panel attrition

Pre-COOL is designed as a panel study. In this section we report the size of panel attrition and which factors predict this attrition. We first look at the predictors of attrition that are observable in the first wave. In a second step we use the register data of all children (also those who left the sample) to track down the occurrence of events such as moving or changes in family structure, and investigate to what extent these events can help explain the attrition in the panel.

Table 2 Approximately half of the children from the initial sample are observed throughout the whole panel

	Wave 3	Wave 4	Wave 5	Complete Panel
N observations in panel	1780 (64%)	1866 (67%)	1544 (55%)	1269 (46%)
N observations not in panel	1005 (36%)	919 (33%)	1241 (45%)	1516 (54%)
Total N observations in wave 1 & 2	2785	2785	2785	2785

Table 2 shows that approximately half of the total sample of children is surveyed in all later waves and can therefore be used for the panel dataset. To learn more about the attrition we estimate the probability of children to be part of measurement waves 3-5 or to be in all of them (complete panel), given that they are either in waves 1 or 2. We take both waves 1 and 2 as the base of the sample, since many children started with preschool only at age 3 and thus show up in wave 2 for the first time. All the control variables in the probit model below are taken from the second wave of pre-COOL (or from the first wave if missing in the second). We only include children for whom we have register data for the whole period.

The results in Table 3 show that children with a migration background are more likely to drop out of the panel structure. If children start out living in a household with a single parent, they are also less likely to still be in the sample in later waves.

Table 3 Children with a migration background are less likely to stay in the panel (panel attrition)

VARIABLES	Wave 3	Wave 4	Wave 5	Complete Panel
Female	0.014	0.021	0.027	0.041**
Migration: Western	-0.042	-0.078**	-0.114***	-0.081**
Migration: non-western second generation	-0.043*	-0.025	-0.048**	-0.058**
Migration: non-western first generation	-0.244**	-0.328***	-0.201*	-0.292***
Household: not married with children	-0.014	-0.010	0.028	0.008
Household: single parent	-0.132***	-0.131***	-0.056*	-0.100***
Parental education: unknown	0.025	-0.021	-0.053	-0.011
Parental education: maximally mbo2	0.042	-0.024	-0.018	0.025
Parental education: mbo 3-4/havo/vwo	0.041*	0.011	0.010	0.018
Percentile household income	0.000	0.000	0.001***	0.001
Observations	2,769	2,769	2,769	2,769

*** significant at p<0.01, ** significant at p<0.05, * significant at p<0.1

Note: Native Dutch is the base-group in migration, the base-group for household is married with children, and hbo/wo is the base-group for parental education.

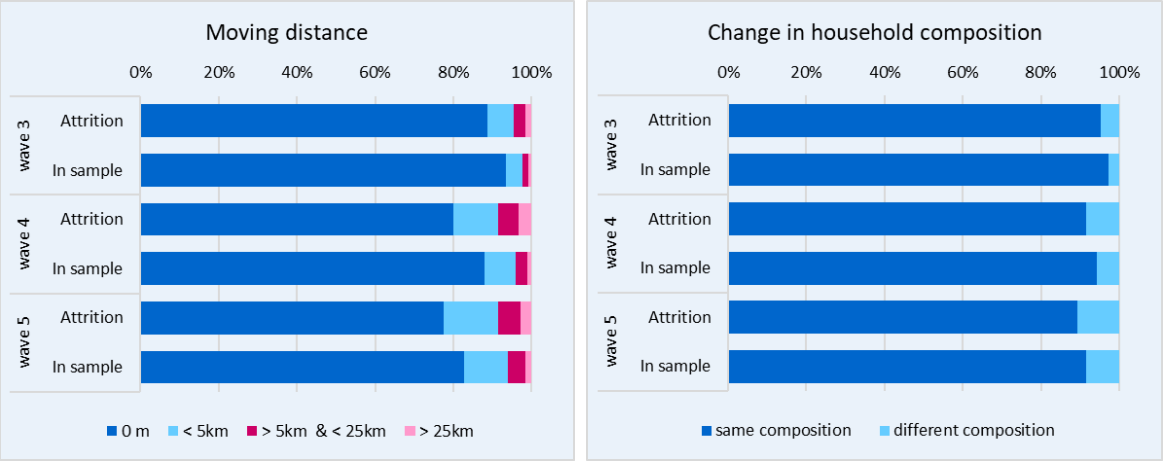
2.4.1 Characteristics and reasons of drop out

The pre-COOL study cannot capture events that take place after the last wave in which children participate in the survey. We therefore make use of the register data, to look at characteristics of the children in the later waves, also for those who are no longer in the panel. This allows us to learn more about the

reasons behind the drop-outs that are based on events that happen after the first measurement and are thus not measured for those who drop out (in the pre-COOL study). In particular we look into two events that are likely to have an impact on whether the child remains in the sample for later measurements or not. A first indicator captures whether a family moved in the period between the second wave of pre-COOL and the later wave, and how far they moved. The second indicator captures if the child still lives with the same parent(s) as in the second measurement wave.

We see that children who are no longer in the sample are more likely to have relocated, and further away. However, since the share of relocations in all groups is rather small, relocations do not explain a large part of the attrition. The left panel of Figure 3 provides a graphical overview of this difference between the children who stay in the sample in later waves versus those who dropped out.

Figure 2 Children in the attrition group are only slightly more likely to move or experience a change in family structure



Children in the attrition group also experience changes in their family structure more frequently. A change in family structure means that at least one caregiver moves in or out of the household. This can happen if the parents separate and one moves out, or if the parent with whom the child lives moves in with a new partner. The difference between children that are in the panel and that are in the attrition group is not very large. Hence, changes in family structure does not provide an explanation for a big share of the attrition form the sample. The right panel of Figure 3 provides a graphical illustration for the difference. It displays whether children are living with the same parent(s) as they did in the second wave of pre-COOL.

2.5 Summary pre-COOL

The pre-COOL study allows to follow a group of preschool children from age two or three to six, when they enter primary education. The study is designed as a panel study and manages to track a sizeable group of children (1269 children, which is approximately 46% of the initial sample) through their early childhood. The attrition from the panel is however non-random, and the use of panel attrition weights at the very least is advised when using the panel structure of the dataset.

Events, such as family relocation or changes in the family structure are related with the probability to drop out of the sample. However, they explain only a small part of the total attrition, and many of the children who drop out of the sample remain living at the same address, and with the same caregivers.

The sample of children that is surveyed for the pre-COOL study is not representative for the reference group of young children of the same age in the Netherlands. Children with a migration background and from families with lower socio-economic status are oversampled. Using population weights is therefore important when using this dataset to learn more about the development of Dutch preschool children.

3 COOL

3.1 Data in brief

The COOL study focuses on children in primary (PO) and secondary (VO) education. The study measures cognitive and non-cognitive skills of school-children in 3 waves, in intervals of 3 years each. COOL was started in 2007 with 4 groups: Group two (6 years old), group five (9 years old), group eight (12 years old) and VO-class 3 (15 years old).[3] In the second wave (2010) a new cohort of children in group two was sampled, while the children who were in group two during the first wave are now sampled in group five. However, while the study aims at testing the same children again in the follow-up waves (2010 and 2013), many schools from the first wave did not participate in later waves, and therefore also new schools were sampled.[4]

The data provides both a panel structure for specific cohorts as well as repeated sampling of the selected age groups in each wave. For the representability study we pool all students of a specific grade across the measurement waves. In the attrition analysis we look at two cohorts of children whom we can observe over the complete span of the study: cohort 1) children who are in group two in 2007 and in group eight in 2013; and cohort 2) children who are in group five in year 2007 and in VO-class three in year 2013.

3.2 Link with CBS register data

Most of the pupils who participate in the COOL cohort study can be linked to CBS register data. However, for 6.6% of all observations in the relevant COOL groups (see above) such a match is not possible. In particular for younger pupils (groups 2 and 5), and in the first measurement wave (2007) the linking fails more frequently.[5]
Using the data that is provided by the cohort study we compare those students who are matched to those who cannot be linked to register data. We find that, as in pre-COOL, children from low SES background are less likely to be matched. We also find a significant difference between the test scores (math, language, as well as work attitude) between the matched and unmatched children, with the latter performing worse.

[3] COOL also samples students in higher grades of VO, which we will however not consider in this document. We discuss only data that is used in the notitie "Ongelijkheid van het jonge kind", which makes use of the data until VO-class 3 (age 15).
[4] A detailed description of the study design and surveys, as well as further literature references can be found at http://www.cool5-18.nl/.
[5] By groups/classes (averaged over all waves) the shares of failing to match are 8% for group two and group five, 6% for group eight and 5% for VO-class three.

3.3 Sample selection

3.3.1 Representativeness

The COOL study is designed as a representative sample of Dutch pupils. However, the representative sample is augmented with an additional sample from schools with a disproportionally large group of students from disadvantaged backgrounds. The additional sample allows to analyze children from disadvantaged backgrounds (as defined by the OAB-weights) in more detail. In this document we use the complete and thus less representative raw sample, since it allows more precise insight in specific groups, such as children with a migration background. In Table 4 we compare this augmented COOL sample to the reference population by grade.[6] The reference population is comprised of all children who are registered in the respective primary or secondary school grade in the Netherlands, in either 2007, 2010 or 2013. In the basic demographic characteristics – gender and age – the COOL samples correspond closely to the representative population.

For the variables of the socioeconomic background the (augmented) COOL sample deviates from the reference population: While migrants are oversampled (by design) in the COOL samples of primary education, the opposite is true for the sample of students from secondary education, i.e. VO-class three, where there is no augmented sample. We find a similar pattern for parental education, children from lower educated parents are oversampled in the COOL sample within primary school, while the children of the COOL VO group are more likely to have higher educated parents than the average. Also in household income we see diverging patterns, with the PO groups coming more frequently from lower income households than the average, while the VO group oversamples students from high income households. Turning to the area in which the children live, we see that the PO cohorts are more frequently living in highly urbanized regions, whereas students in the VO sample are more likely to live in an area of low urbanization.

Table 4 The COOL sample differs from the reference population in key variables, such as migration background

	Group 2		Group 5		Group 8		VO-Class 3	
VARIABLES	Reference population	COOL	Reference population	COOL	Reference population	COOL	Reference population	COOL
Female	0.486	0.479	0.494	0.498	0.494	0.506	0.494	0.502
Migration: Native	0.763	0.681	0.773	0.688	0.777	0.699	0.785	0.815
Migration: Western	0.064	0.061	0.061	0.060	0.061	0.057	0.061	0.058
Migration: non-western second generation	0.163	0.246	0.152	0.236	0.141	0.219	0.127	0.107
Migration: non-western first generation	0.010	0.011	0.014	0.016	0.022	0.025	0.026	0.021
Parental education: unknown	0.171	0.171	0.213	0.209	0.247	0.238	0.278	0.276
Parental education: maximally mbo2	0.204	0.289	0.206	0.285	0.218	0.288	0.223	0.195
Parental education: mbo3-4/havo/vwo	0.260	0.269	0.256	0.258	0.256	0.256	0.256	0.262
Parental education: hbo/wo	0.365	0.271	0.326	0.248	0.278	0.218	0.242	0.266
Household income quintile: unknown	0.011	0.009	0.011	0.009	0.012	0.008	0.010	0.007

[6] We report the group averages over the three sample years, since we use this specification also for our main results.

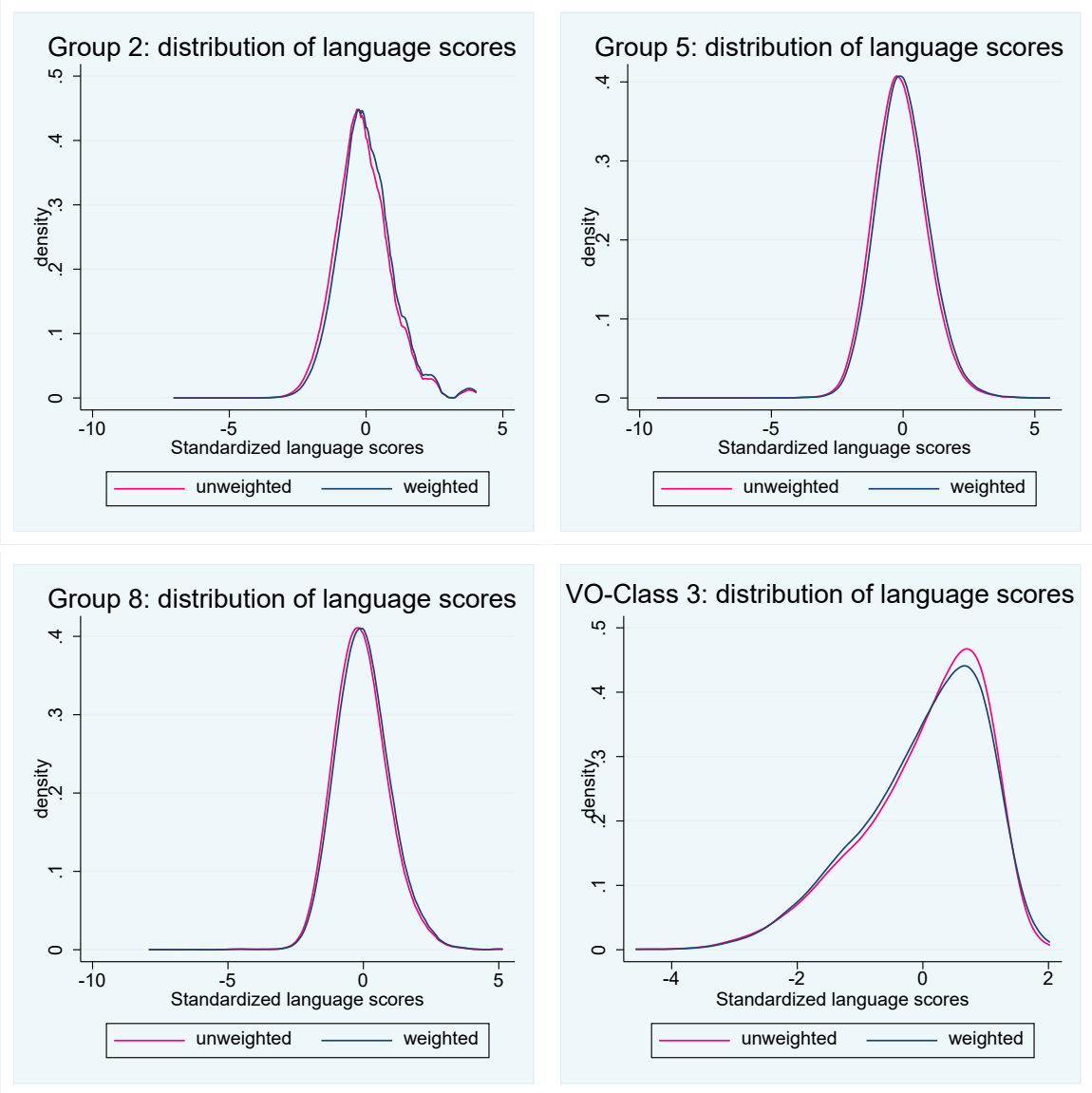
Household income quintile: lowest	0.229	0.300	0.216	0.280	0.200	0.256	0.164	0.149
Household income quintile: second	0.208	0.234	0.200	0.219	0.191	0.213	0.181	0.178
Household income quintile: middle	0.223	0.211	0.220	0.215	0.221	0.219	0.229	0.235
Household income quintile: fourth	0.183	0.148	0.192	0.164	0.205	0.179	0.229	0.241
Household income quintile: highest	0.146	0.097	0.162	0.113	0.171	0.125	0.187	0.191
Type of household: unknown	0.011	0.010	0.010	0.010	0.013	0.009	0.012	0.009
Type of household: couple with children (married)	0.184	0.170	0.139	0.129	0.107	0.103	0.081	0.081
Type of household: couple with children (married)	0.690	0.684	0.717	0.710	0.727	0.719	0.735	0.756
Type of household: single parent	0.115	0.136	0.134	0.151	0.153	0.169	0.171	0.154
Age	5.8	5.8	9.0	8.9	12.0	12.0	15.1	15.0
Urbanization : unknown	0.005	0.003	0.004	0.003	0.006	0.002	0.003	0.002
Urbanization : high	0.452	0.474	0.434	0.464	0.423	0.453	0.422	0.364
Urbanization : medium	0.206	0.183	0.207	0.176	0.208	0.180	0.210	0.177
Urbanization : low	0.337	0.340	0.354	0.357	0.363	0.365	0.365	0.457
Number of obs.	598272	3203	586543	32648	589644	32146	591846	44054

3.3.2 Population weights

The section above shows that the samples deviate from the reference population in key background characteristics, and that not all of the groups deviate in the same manner. If we combine the samples in a synthetic panel without any correction, we are at risk of misinterpreting sampling biases as actual results. Specifically, in our study on the development of skills (notitie “Ongelijkheid van het jonge kind”) we standardize test-scores within age groups. Without controlling for changes in the compositions of the samples, the students would be measured on different standards, depending on whether better or worse students are oversampled.

The bias in sampling has a small, but non-negligible impact on the distribution of cognitive skills that are measured. Figure 4 shows the density plots of the language scores for each of the four tested age groups (grades). The plots reveal that in PO the unweighted distributions of the language scores (observed) are somewhat lower than the weighted distributions of these scores (corrected), while in VO this difference is reversed. This is in line with the different group composition in PO and VO that can also be seen in the background characteristics in the chapter above. We find similar differences in the distributions for math, but no differences in the distribution of work attitude (werkhouding).

Figure 3 Distribution of language test scores in the weighted and unweighted COOL sample



3.4 Panel attrition

Unlike the pre-COOL study, the COOL study does not follow individual students, but rather surveys specific grades of schools. This leads to a panel data set for only those students, who progress to through school on the regular time scheme, without repeating or skipping grades.[7] The targeted grades are group two, five, eight and VO-class three. We thus have two potential cohorts: cohort 1) children start in group two in 2007 and are in group eight in year 2013; and cohort 2) children start in group five in 2007 and are in VO-class three in 2013. The second cohort transitions from primary school to secondary school during the measurement period. This transition comes with a change in school. The COOL study samples classes of secondary schools which are likely to receive many children from the previously sampled primary school

[7] This issue is also discussed in detail in: Zittenblijven en verwijzing in het basisonderwijs. Onderzoek naar uitvallers uit de tweede meting van het cohortonderzoek COOL5-18. Jaap Roeleveld, Hermann Vierke & Lia Mulder - Nijmegen: ITS / Amsterdam: Kohnstamm Instituut

classes. However, many children are lost during this step, while many new children are added to the sample. Table 5 shows per cohort how many children from the initial measurement are in the panel.

Table 5 The panel data sets include only a small share of all students from the respective initial samples

	Cohort 1 (start in group 2)			Cohort 2 (start in group 5)		
	Group 5 Wave 2 (2010)	Group 8 Wave 2 (2013)	Complete Panel	Group 8 Wave 2 (2010)	VO-Class 3 Wave 2 (2013)	Complete Panel
N observations in panel	5282 (44%)	2550 (21%)	2324 (19%)	5755 (52%)	1551 (14%)	928 (8%)
N observations not in panel	6858 (56%)	9590 (79%)	9816 (81%)	5371 (48%)	9575 (86%)	10198 (92%)
Total N observations in initial sample (2007)	12140	12140	12140	11126	11126	11126

The panel attrition is large and systematically related to the outcomes (e.g. through grade repetition). Therefore using the panel structure is problematic. We discuss in this section the small subsample that is available for panel sample, and which selection mechanisms lead to this group. Since a large share of the attrition is due to schools dropping out of the sample or students repeating a grade, we start first with the investigation of events that happen after the first measurement. In a second step we show the attrition analysis where we estimate relation between the probability to stay in the sample and individual characteristics.

3.4.1 Characteristics and reasons of drop out

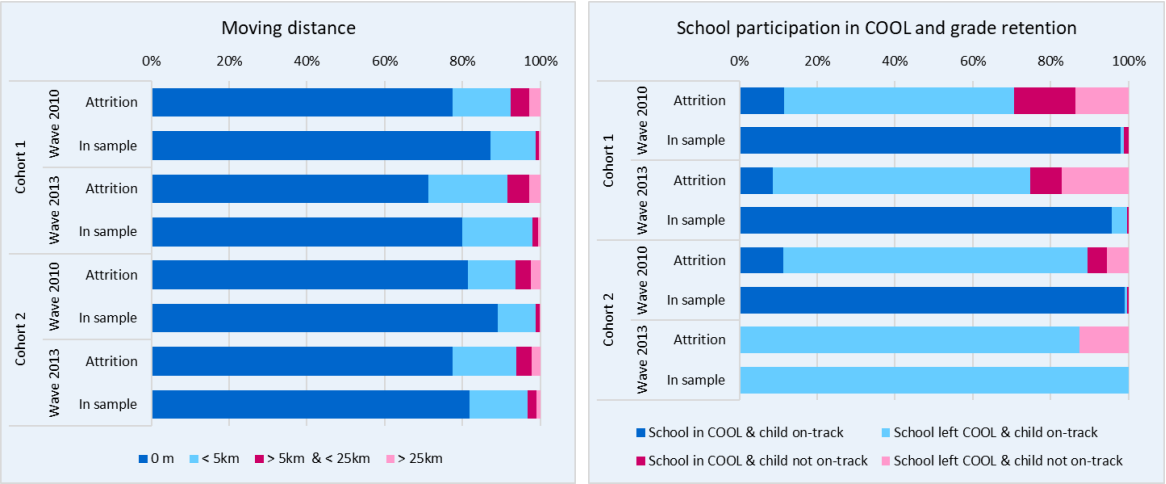
We investigate three major reasons for dropping out of the sample: 1) Students whose families relocate between the first and the second or third wave drop out of the sample if their new school is not part of the COOL study. 2) Many schools have not participated in the complete COOL study, but rather in one or two waves. If a school no longer participates, all the students of this school drop out of the COOL study. 3) If a student repeats or skips a grade, they will no longer be in a class that is surveyed, even if their school still participates in the following waves of the COOL study.

Children who dropped out are more likely to have moved, and in particular to have moved further away. This difference is illustrated graphically in the left panel of Figure 3.2. While the pattern aligns with our expectations, moving explains at most 30% of the attrition, much less if we take into account that not all students who moved also drop out of the sample.

School dropout is an important factor for attrition. The right panel of Figure 3.2 combines both school related factors (school dropout and grade repetition). We see that in the second wave more than 70% of the students in Cohort 1 went to a school which discontinued their participation in the COOL study. In the third wave more than 80% of the loss of students is due to schools not participating in this wave. In the second cohort all the schools are discontinued in the second wave, when the students change from primary to secondary education. The secondary schools in the COOL study are selected in such a way, that they are likely to be receiving schools for COOL students. This allows to follow some students to secondary education in a panel setting, even though those students are no longer in the original COOL participating school.

Grade repetition is a further important reason for attrition. The graph also shows that grade repetition is especially prevalent between groups two and five and thus influences the attrition for the first cohort particularly strongly.

Figure 4 Much of the panel attrition in COOL is due to school dropout or grade retention



3.4.2 Panel structure and predictors of attrition

A large share of the attrition is attributable to decisions made by schools, but grade repetition also plays a role. If school drop-outs are the driving force behind the attrition, this could mean that we should expect mostly random attrition among students. However, if the schools which drop out are not random, but rather selective on the basis of the average SES of their pupils or migration background, the panel sample will get more selective as well. Additionally, since grade repetition is a reason for dropping out of the panel, worse performing students are less likely to be in the panel data set. We investigate this selective attrition by estimating the characteristics of the children in wave 1 on their probability be part of measurement waves 2 and 3, or the complete panel.

The results in Table 6 show that the attrition is not random. Within the first cohort – in which grade repetition is an important reason for attrition – girls are more likely to stay in the COOL sample than boys. Children with a migration background have a higher chance to stay in the panel within the primary education, and a lower chance in the secondary education (wave 3 of the second cohort). Unlike in the pre-COOL sample, the urbanization of the place of residence (measured in the first wave) matters for sample attrition, children who live in highly urbanized regions are more likely to drop out of the sample.

Table 6 Probability to be in (panel) sample in measurement waves 2 and 3 is not random

	Cohort 1			Cohort2		
VARIABLES	**Wave 2**	**Wave 3**	**All waves**	**Wave 2**	**Wave 3**	**All waves**
Female	0.030***	0.018**	0.019***	0.008	0.007	-0.000
Migration: Western	0.010	-0.031**	-0.022	-0.004	0.004	-0.001
Migration: non-western second generation	0.046***	0.049***	0.058***	0.042***	-0.037***	-0.010
Migration: non-western first generation	-0.021	-0.043	-0.019	-0.031	-0.045**	-0.023
Household: not married with children	-0.032**	-0.016	-0.012	-0.021	-0.018*	-0.005
Household: single parent	-0.069***	-0.045***	-0.044***	-0.046***	-0.026***	-0.018**
Household: other	-0.039	-0.065	-0.047	-0.002	-0.023	-0.031
Parental education: unknown	0.017	0.012	0.009	0.042***	-0.029***	-0.005
Parental education: maximally mbo2	-0.034**	-0.028**	-0.031***	0.025*	-0.039***	-0.011
Parental education: mbo 3-4/havo/vwo	-0.001	-0.010	-0.010	0.018	-0.028***	-0.004
Percentile household income	0.000	0.000*	0.000*	0.000*	0.000*	0.000
Urbanization: medium	0.042***	0.062***	0.042***	0.061***	0.049***	0.026***
Urbanization: low	0.079***	0.031***	0.035***	0.072***	0.033***	0.039***
Urbanization: missing	-0.065	-0.073	-0.060	-0.039	0.025	-0.014
Observations	12,140	12,140	12,140	11,126	11,126	11,126

*** significant at p<0.01, ** significant at p<0.05, * significant at p<0.1

Note: Native Dutch is the base-group in migration, the base-group for household is married with children, and hbo/wo is the base-group for parental education.

3.5 Summary COOL

The COOL study surveys cohorts of Dutch pupils repeatedly, allowing to analyze the development of skills over time. Using all available data points (including the augmented sample), we find that certain groups of children are oversampled. While children from low SES families have a higher probability to be part of the study in primary school, the exact opposite is true for the pupils in secondary education. Using population weights is therefore necessary for the full COOL dataset, in particular if the cohorts are combined in a synthetic panel.

The COOL data not well suited for panel studies. The design of the study is to follow classes, not individual students. As consequence of this design, panel attrition is directly linked to the outcome variables, i.e. children who perform poorly are at risk of repeating a grade and thus more likely to drop out from the panel. Readjusting the sample with panel attrition weights does not solve the bias that results from this selective attrition. It is hence inadvisable to use the COOL dataset for longitudinal studies that focus on individual learning improvements. This issue is discussed by Kohnstamm and the COOL dataset accordingly emphasizes the cohort, rather than the panel dimension.